

Syntaktická rovina a syntaktické funkce

1. Úvod

Nejprve uvedeme stručný přehled rovin v celém systému. Celý systém obsahuje tři roviny, čili každá věta je reprezentována na těchto rovinách:

- a) slovní rovina (SLOVR)
- b) morfologická rovina (MORFR)
- c) **syntaktická rovina (SYNTR)**

2. Slovní rovina (SLOVR)

Na slovní rovině SLOVR jsou jednotlivá ortografická tokenizovaná slova, například *rozhodl se* jako dvě slova. Slova jako *viděls, očs, včeras, vltavotýnský, hornolužický, společenskopolitický* jsou na této rovině jako taková. Každé interpunkční znaménko tvoří zvláštní token. Elipsa (tři tečky) je zachycena jako jeden interpunkční token.

Na rovině SLOVR není žádná anotace, jen výsledek tokenizace a větné segmentace.

3. Morfologická rovina (MORFR)

Každému prvku na rovině MORFR je přiřazeno jediné lemma a jediná morfologická značka jako výsledek automatické morfologické analýzy a disambiguace.

Příklad 1

Mějme konstrukci *rozhodl se* ve větě:

Rozhodl se pracovat.

Na MORFR odpovídají spojení *rozhodl se* dva uzly:

- uzel pro *rozhodl*
- uzel pro *se*.

Uzel pro *rozhodl* obsahuje na MORFR

- i. **jediné** lemma *rozhodnout*
- ii. **jedinou** morfologickou značku VpMS---3R-AA--1P nebo VpIS---3R-AA--1P (podle disambiguace v konkrétní větě)

Uzel pro nepředložkové *se* obsahuje na MORFR

- i. **jediné** lemma *se*
- ii. **jedinou** morfologickou značku P7---4-----

4. Syntaktická rovina (SYNTR)

Na rovině SYNTR je struktura věty zachycena bezprostředněsložkovým stromem, v němž jsou *terminální* a *neterminální* uzly. Každý terminální uzel má **funkci**; každý neterminální uzel je jistého **typu** a s výjimkou řídicího uzlu celé věty má též **funkci**.

V bezprostředněsložkové struktuře na rovině SYNTR nejsou žádné zvláštní uzly pro lexikálně nevyjádřené elementy věty, například pro subjekt finitního predikátu nebo pro subjekt infinitivu.

V bezprostředněsložkové struktuře na rovině SYNTR nejsou ani žádné zvláštní uzly pro interpunkci členící větu nebo vyjadřující otázku, zvolání apod. (čárky, tečky, vykřičníky, otazníky). Tyto elementy figurují jen na rovině grafémické a morfologické.

4.1 Typy

Typy neterminálních uzlů na nejvyšší typové úrovni jsou dva:

Typ A. TypeUnheaded

Typ B. TypeHeaded

Typ A. TypeUnheaded

Typ **TypeUnheaded** má dále tyto podtypy:

- **Coord**: koordinace
- **Apos**: apozice
- **MWE** (multi-word expression): dále nestrukturovaný víceslovný výraz

tj. schematicky:

TypeUnheaded [Coord|Apos|MWE]

Uzel typu **TypeUnheaded** může fakultativně obsahovat morfosyntaktické kategorie. Jednotlivé složky (tj. dcery) nadsložky typu **Coord** | **Apos** | **MWE** se k této složce připojují jednou z těchto speciálních funkcí:

- **CoMemb**
- **ApMemb**
- **MWEMemb**

(viz též níže v přehledu funkcí). Význam koordinace a apozice je jasný, podtyp MWE vystihuje jakýkoli dále nestrukturovaný víceslovný výraz, např. frazém, kolokaci, případně i adordinační spojení. Je-li součástí uvedených typů spojka, je vyjádřena jako zvláštní složka, již je typicky přiřazena funkce

CoConj pro spojku v koordinované konstrukci. Není-li spojka součástí koordinace (její součástí jsou např. pouze čárky), není součástí vyjádření koordinace žádná složka označená funkcí **CoConj**.

Příklad 2

Spojení: *otec a matka nebo sestra* se vyjádří buď takto:

[[*otec-CoMemb matka-CoMemb a-CoConj*] *sestra-CoMemb nebo-CoConj*]-Fun

nebo

[*otec-CoMemb [matka-CoMemb sestra-CoMemb nebo-CoConj]-CoMemb a-CoConj*]-Fun

kde odlišná vyjádření zachycují odlišnou těsnost konjunktů. *Fun* je proměnná označující funkci celé koordinace ve větě, např. funkci objektu (Obj) ve větě (obj), resp. subjektu (Subj) ve větě (subj):

(obj) *Vidím otce a matku nebo sestru.*

(subj) *Otec a matka nebo sestra odešli.*

Příklad 3

Víceslovné výrazy s jedinou funkcí, například víceslovné spojky, jsou vyjádřeny pomocí funkce **MWEMemb**. Například víceslovná spojka *i když* je vyjádřena takto:

[*i-MWEMemb když-MWEMemb*]-Fun

kde Fun je prakticky vždy rovno funkci SurfHead (viz níže).

Podobně se vyjádří i elementy frazému.

4.2 Funkce

Podobně jako uzly typu **TypeUnheaded** i uzly typu **TypeHeaded** mají složky (tj. dcery) a každá z těchto dcer má právě 1 funkci (až na výjimky označené jako * – viz níže) z tohoto repertoáru:

Advb – adverbiale

AdvbAuxZ – zdůrazňovací slovo, které zdůrazňuje slovo jiného slovního druhu než syntaktické substantivum

Attr – atribut

AttrAuxZ – zdůrazňovací slovo, které zdůrazňuje syntaktické substantivum

Deagent – reflexivum v deagentivním významu

DeepHead – hloubková hlava, může mít jako alias funkci **NomPred** nebo **Advb**

***Head** – SurfHead & DeepHead (= konjunkce obou funkcí)

NomPred – jmenný přísudek

Obj – objekt

ReflTant – reflexivum tantum

Subj – subjekt

SurfHead – povrchová hlava

VbAttr – doplněk tvořící složku s jiným než slovesným tvarem

VbAttrV – doplněk tvořící složku se slovesným tvarem

AuxO – nadbytečný element (podle PDT)

AuxY – „zbytkové“ příslovce nebo částice (obvykle však vyjádřené jako SurfHead) (podle PDT)

AuxG – grafický symbol s jiným významem ve větě než organizačním (podle PDT)

ExD – ExDependent (podle PDT)

ApMemb – člen apozice

CoMemb – člen koordinace

CoConj – spojka v koordinaci

MWEMemb – člen víceslovného výrazu

Uvedené funkce mohou mít obecně asociovanou množinu aliasů. Reálně však může mít aliasy pouze funkce **DeepHead**, a to:

NomPred – označuje nominální predikát včetně pasíva ve strukturách se sponou vyjádřenou tvarem slovesa *být/bývat* (*Chlapec byl hodný/přistižen.*)

Advb – označuje adverbíale v konstrukcích s tvarem slovesa *být/bývat* (*Byl jsem doma.*)

4.2.1 Funkce SurfHead a DeepHead

a. Všechny tvary pomocného slovesa *být/bývat* (pro tvoření minulého a předminulého času, minulého a předminulého kondicionálu, opisného futura a opisného pasíva) včetně spony mají funkci **SurfHead**. Totéž platí i o kondicionálové částici *by*. Jenom existenční *být* má funkci **Head**. Je to proto, aby se ze struktury věty na SYNTR dala snadno odvodit struktura povrchová (řídícím prvkem takové struktury je SurfHead) či hloubková (řídícím prvkem takové struktury je DeepHead). Podstrom se složkami SurfHead a DeepHead je vždy binární.

Například ve větě:

Čeněk je hlupák/hloupý/zatčen.

má spona *je* funkci **SurfHead** a nominální predikát *hlupák/hloupý/zatčen* funkci **DeepHead** alias **NomPred**.

Tvar sponového *být* je jedinou sestrou příslušného nominálního predikátu. Zbývající doplnění složeného predikátu (spona-SurfHead, nominální_predikát-NomPred) jsou sesterskými uzly otcovského uzlu tohoto složeného predikátu. Tento přístup umožní nenarušit binaritu podstromu: (uzelx-SurfHead uzely-DeepHead). Funkce SurfHead a DeepHead se tedy mohou vyskytnout jen spolu v binárním stromě.

b. Zástupné, dispoziční *být/bývat*:

Tento typ lze exemplifikovat těmito příklady:

Táta je doma.

Táta byl na zahradě.

Boty jsou mi dobře.

Spona je tu sestrou nominálního predikátu, zbývající elementy závisejí na dvojici [spona, nominální predikát], a jsou tedy sestrami neterminálu reprezentujícího tuto dvojici.

Příklad 4

Věta:

Boty jsou mi dobře.

má strukturu:

[*boty*-Subj [[*jsou*-SurfHead *dobře*-DeepHead]-Head *mi*-Obj]-Head]

Funkce **DeepHead** zde má ovšem alias **NomPred**.

Dvojici (spona, nominální predikát) zde tvoří: (*jsou, dobře*), nikoli (*jsou, mi*)!

Věta:

Táta je na zahradě s Pavlem.

má strukturu:

[*táta*-Subj [[*je*-SurfHead [*na*-SurfHead *zahradě*-DeepHead]-DeepHead]-Head [*s*-SurfHead *Pavlem*-DeepHead]-Obj]-Head]

Funkce **DeepHead** zde má alias **Advb**.

Čili [*je, na zahradě*] chápeme jako jeden složený predikát, kde *je* je spona a [*na zahradě*] má funkci **DeepHead** alias **Advb**.

c. Stav a děj

Strukturně se na rozdíl od PDT nerozlišuje *stav* a *děj*. Pomocné sloveso se vždy chápe jako spona (**SurfHead**) a jmenný přísudek (pasivní participium/syntaktické adjektivum/syntaktické substantivum) jako **DeepHead**_alias_**NomPred**.

4.2.2 Pomocná, modální a fázová slovesa

a. Mít/mívat, dostat/dostávat jako pomocná slovesa

V rezultativních konstrukcích typu *Mám uvařeno* má pomocné sloveso *mám* na rozdíl od spony přisouzenou funkci **Head** a *uvařeno* funkci **VbAttr**.

Příklad 5

Věta:

Mám uvařeno.

má takovouto strukturu:

[*mám*-Head *uvařeno*-VbAttr]

b. Modální a fázová slovesa

Tato slovesa mají funkci **Head**.

Příklad 6

Ve větě:

Začal/Chci pracovat.

má *začít/chci* funkci *Head* a *pracovat* funkci *Obj*:

[*začal*-Head *pracovat*-Obj]

4.2.3 Reflexivita

K dispozici jsou 3 funkce:

- **Obj**
- **ReflTant**
- **Deagent**

Níže uvádíme podtypy reflexivity v našem pojetí.

a. Reflexiva tantum

Reflexivní částice *se/si* má funkci **ReflTant**.

Příklady 7

Zasmála se.

Libí se mi to.

To se nedá oddiskutovat.

Usmyslil si odejít.

Všimla si toho.

Muž chvějící se hrůzou hrůzoucí.

Přízpůsobení se světu po 2. vatikánském koncilu

Zde sloveso / adjektivum / substantivum se sestrou reflexiva *se/si*; sloveso / adjektivum / substantivum má funkci **Head** a příslušné reflexivum funkci **ReflTant**.

b. Poloreflexiva tantum

I zde má reflexivní částice *se/si* funkci **ReflTant**.

Příklady 8

Rozhodl se to udělat.

Posadil se na židli.

Vlívá se do moře.

Vlívání se do moře.

c. Reflexivum objektové

Reflexivní zájmeno *se/si* má funkci **Obj**:

Příklad 9

Myji se.

d. Reflexivum mediopasivní

Reflexivní zájmeno *se/si* zde má funkci **Obj** nebo **Advb**.

Příklady 10

Myslím si to.

Přečti si to.

Tak si zavaž tkaničku.

Postupte si do vozu.

e. Reflexivum pasivní v deagentivní a ergativní konstrukci

Reflexivní částice *se/si* zde má funkci **Deagent**.

Příklady 11

To se nedělá.

Tady se nepracuje.

To se ti to pracuje.

Nechce se mi.

f. Reflexivum vyjadřující dispoziční modalitu

Reflexivní částice *se/si* zde má funkci **Deagent**.

Příklady 12

V troubě se dobře peče.

g. Reflexivum reciproční

Reflexivní zájmeno *se/si* zde má **Obj.**

Příklady 13

Koulovali se.

Kočkují se.

Milují se.

Dávají si pusy.

Macek s Rathem se fackují/fackuje.

Nezavádíme zvláštní funkci pro reciproční reflexivitu, protože reciproké konstrukce se zatím automatickými nástroji rozpoznat nedají. Vše, co má životný objekt, může být totiž reciproční.

Literatura

Prague Dependency Treebank 2.0. ÚFAL MFF UK, Praha. Dostupný z WWW:
<<http://ufal.mff.cuni.cz/pdt2.0>>

Hajič J., J. Panevová, E. Buráňová, Z. Uřešová, A. Bémová, J. Kárník, J. Štěpánek, P. Pajas (1999):
Pražský závislostní korpus – PDT. Anotace na analytické rovině. Návod pro anotátory. ÚFAL MFF UK, Praha.